

# Propensity Score Workshop

Timothy Feeney

# Table of contents I

Preliminaries

What are we aiming to do here?

## Preliminaries

# Teaching Team Introductions

- ▶ Tobias Kurth, Charité, Berlin Germany
- ▶ Catie Weiner, UNC Chapel Hill, USA
- ▶ Timothy Feeney, UNC Chapel Hill, USA

# Outline of the workshop

## 1) Didactics

- ▶ background and theory of basic propensity scores
- ▶ estimands discussion
- ▶ theory of inverse probability weighting
- ▶ theory of standard error (SE) estimation

## 2) Practical Activities

- ▶ Breakout sessions to practice and use what we are teaching
- ▶ Will do propensity score analysis, bootstrapping, m-estimation
- ▶ Will use R for coding. RMarkdown Files have been provided for reference

# Schedule for the Afternoon

## First half

1:10-1:20 Introduction to Clinical Problem

1:20-1:40 First Activity:

- ▶ Descriptive and Crude Analysis

1:40-2:40 Didactics:

- ▶ Overview of causal, estimands, propensity score basics

2:40-3:10 Second Activity:

- ▶ Estimation of ATE and ATT using IPTW

3:10-3:30 Break

# Schedule for the Afternoon

## Second half

3:30-4:10 Didactics:

- ▶ Standard Error estimation using Bootstrapping
- ▶ Standard error estimation using M-estimation

4:10-4:50 Third Activity:

- ▶ Using Bootstrap and M-estimation

4:50-5:00 Questions and Closing Remarks

# Hearing from you

What do you sepcifically want to get out of this workshop?



What are we aiming to do here?

# Research Types

	Description	Prediction	Causal inference
Example of scientific question	How can women aged 60–80 years with stroke history be partitioned in classes defined by their characteristics?	What is the probability of having a stroke next year for women with certain characteristics?	Will starting a statin reduce, on average, the risk of stroke in women with certain characteristics?
Data	<ul style="list-style-type: none"> <li>• Eligibility criteria</li> <li>• Features (symptoms, clinical parameters ...)</li> </ul>	<ul style="list-style-type: none"> <li>• Eligibility criteria</li> <li>• Output (diagnosis of stroke over the next year)</li> <li>• Inputs (age, blood pressure, history of stroke, diabetes at baseline)</li> </ul>	<ul style="list-style-type: none"> <li>• Eligibility criteria</li> <li>• Outcome (diagnosis of stroke over the next year)</li> <li>• Treatment (initiation of statins at baseline)</li> <li>• Confounders</li> <li>• Effect modifiers (optional)</li> </ul>
Examples of analytics	Cluster analysis ...	Regression Decision trees Random forests Support vector machines Neural networks ...	Regression Matching Inverse probability weighting G-formula G-estimation Instrumental variable estimation ...

Figure 1: Three Broad Types of Research—We will focus on the third column

# Propensity Score Estimation-Theory

- ▶ linkage to the g-formula

# Types of Propensity Score Estimation

- logistic regression
- High dimensional PS?
- multiple levels of treatment?
- ML approaches? (brief)

# Uses of Propensity Scores

- Stratification
- Regression
- Matching
  - Time-dependent matching?
  - Two-stage matching?

# Propensity Score Diagnostics

- Distribution of scores
- Overlap
- Trimming?
- Truncation?

# Stabilized Versus Unstabilized

# Standard Error Estimation-Bootstrapping



# Standard Error Estimation

M-Estimation